Use of Novel Algorithms MAJE4 and MACJER-320 for Achieving Confidentiality and Message Authentication in SSL & TLS

Sheena Mathew, K. Poulose Jacob

Abstract—Extensive use of the Internet coupled with the marvelous growth in e-commerce and m-commerce has created a huge demand for information security. The Secure Socket Layer (SSL) protocol is the most widely used security protocol in the Internet which meets this demand. It provides protection against eavesdropping, tampering and forgery. The cryptographic algorithms RC4 and HMAC have been in use for achieving security services like confidentiality and authentication in the SSL. But recent attacks against RC4 and HMAC have raised questions in the confidence on these algorithms. Hence two novel cryptographic algorithms MAJE4 and MACJER-320 have been proposed as substitutes for them. The focus of this work is to demonstrate the performance of these new algorithms and suggest them as dependable alternatives to satisfy the need of security services in SSL. The performance evaluation has been done by using practical implementation method.

Keywords—Confidentiality, HMAC, Integrity, MACJER-320, MAJE4, RC4, Secure Socket Layer

I. INTRODUCTION

WITH the growth of the Internet and digital transmission, many applications need to transmit data to remote applications and computers securely. The four main stated security issues are confidentiality, authentication, integrity and non-repudiation. Confidentiality means, only authorized users can access the information while unauthorized users are to be denied access. Authentication has to guarantee that the user accessing the information is truly the intended person and not a pretender. Integrity has to ensure that the received information is same as the transmitted information while being modified by others during transmission. Non-repudiation guarantees that senders and receivers have undeniably transmitted or received information, respectively. These four mentioned issues are interdependent and must therefore be addressed simultaneously in the design of security systems. SSL protocol has been universally accepted in the World Wide Web for authenticated and encrypted communication between clients and servers.

The SSL protocol was originally developed by Netscape, its version 1.0 was never publicly released; version 2.0 was released in 1994 but contained a number of security flaws which ultimately led to the design of version 3.0 which was released in 1996[1]. At present, SSL is widely deployed in many intranets as well as over the public Internet and has become the de facto standard for transport layer security. Recently, the Internet Engineering Task Force (IETF) has started an effort to standardize SSL as an IETF standard under the name of Transport Layer Security (TLS) protocol [2]. The few real world, practical applications of SSL & TLS are [3] client server systems, financial systems, information systems to create remote access and administration applications, travel industry to create online reservation systems and secure information transfer, etc. Visa, MasterCard, American Express and many leading financial institutions have endorsed SSL for commerce over the Internet. Some early implementations of SSL used 40-bit symmetric keys because of US government restrictions on the export of cryptographic technology. The 40-bit key size limitation has mostly gone away and modern implementations use 128-bit (or longer) keys for symmetric key ciphers.

One of the reasons that SSL has outgrown other transport and application layer security protocols such as SSH, SET, and SMIME in terms of deployment is that it is application protocol independent [4]. Conceptually, any application that runs over TCP can also run over SSL. There are many examples of applications such as TELNET and FTP running transparently over SSL. However, SSL is most widely used as the secure transport layer below HTTP. A large number of e-commerce sites dealing with private and sensitive information use SSL as the secure transport layer. This number is expected to grow, as more and more businesses and users embrace electronic commerce. As security becomes an integral feature of Internet applications and the use of SSL rises, its impact on the performance of the servers as well as the clients is going to be increasingly important. Browsers like Netscape Navigator and Internet Explorer can access SSL enabled web pages by using URLs that start with ‘https’ instead of ‘http’.

The main objectives for SSL are:
1. Authenticating the client and server to each other.
2. Ensuring data integrity
II. MOTIVATION

In applications using SSL, the confidentiality of information is ensured using strong encryption algorithms. For very fast encryption and decryption of data for transmission after an SSL connection has been established, RC4 is the preferred algorithm. Similarly HMAC-SHA-1 has been recommended for message authentication in several network security protocols. The key reasons behind this are the free availability, the flexibility of chaining the hash function and the reasonable speed, among others. Even though RC4 and HMAC-SHA-1 are the most widely used ciphers of secure web applications, the strength of RC4 [5] and SHA-1 [6] has been called into question as a result of recent findings. Hence it is required to have proven and new methods to meet the future requirements. The analysis of novel cryptographic algorithms MAJE4 [7] and MACJER-320 and its performance in comparison with the popular RC4 and HMAC-SHA1 have been done in this context and the novel algorithms MAJE4 and MACJER-320 have been proposed as alternatives.

III. DESCRIPTION OF EXISTING ALGORITHMS

This section presents an overview of the cryptographic algorithms RC4 and HMAC-SHA-1 used in SSL.

A. RC4

RC4 is a variable key-size stream cipher developed in 1987 by Ron Rivest for RSA Data Security, Inc. The RC4 stream cipher has two phases, the key set-up and the keystream generation. Both phases must be performed for every new key. The algorithm is based on the use of a random permutation. A variable length key K, of size 1 to 256 bytes is used to initialize a 256-byte state vector S, with elements $S_0, S_1, \ldots, S_{255}$.

Initially the entries of S are set to the values 0 to 255 in ascending order. A temporary vector T, is also created. For a key of length keylen bytes, the first keylen elements of T are copied from K, and then K is repeated as many times as necessary to fill out T. Next, we use T to produce the initial permutation of S. The pseudo-code for the key setup is as follows:

\[
\begin{align*}
\text{for } i &= 0 \text{ to } 255 \\
S_i &= i \\
T_i &= K[i \mod \text{keylen}] \\
\text{endfor} \\
k &= 0 \\
\text{for } i &= 0 \text{ to } 255 \\
k &= (k + S_i + T_i) \mod 256 \\
\text{Swap}(S_i, S_k) \\
\text{endfor}
\end{align*}
\]

Once S is initialized, the input key is no longer used. The next phase is key stream generation which is described by the pseudo-code as:

\[
\begin{align*}
i &= (i + 1) \mod 256 \\
k &= (k + S_i) \mod 256 \\
\text{Swap}(S_i, S_k) \\
t &= (S_i + S_k) \mod 256 \\
\text{key} &= S_t \\
\text{endloop}
\end{align*}
\]

For encryption, the value key is XORed with the next byte of plaintext. For decryption, the value key is XORed with the next byte of ciphertext.

B. HMAC

The different variables used in the HMAC algorithm are shown in Table I.

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASIC NOTATIONS IN HMAC</td>
</tr>
<tr>
<td>MD - Message digest/ hash function</td>
</tr>
<tr>
<td>M - Input message</td>
</tr>
<tr>
<td>B - Number of bits in each block</td>
</tr>
<tr>
<td>K - Shared symmetric key</td>
</tr>
<tr>
<td>K1 - Transformed key K1</td>
</tr>
<tr>
<td>Ipad - String 00110110 repeated b/8 times</td>
</tr>
<tr>
<td>Opad - String 01011010 repeated b/8 times</td>
</tr>
<tr>
<td>H - Hash code</td>
</tr>
</tbody>
</table>

The step-by-step approach of the HMAC message authentication code is given in Table II.

<table>
<thead>
<tr>
<th>TABLE II</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMAC ALGORITHM</td>
</tr>
<tr>
<td>Step 1: Make the length of K equal to B. Append enough zeros to the left end of K to create a B bit key K1.</td>
</tr>
<tr>
<td>Step 2: XOR K1 with Ipad to produce the B bit block S.</td>
</tr>
<tr>
<td>Step 3: Append M to S. That is the original message is simply appended to the end of S.</td>
</tr>
<tr>
<td>Step 4: Apply the Message digest algorithm/ Hash function to the output of Step 3 to produce hash code H.</td>
</tr>
<tr>
<td>Step 5: XOR K1 with Opad to produce the B bit block S1.</td>
</tr>
<tr>
<td>Step 6: Append the hash code H produced in Step 4 to S1.</td>
</tr>
<tr>
<td>Step 7: Apply the message digest algorithm/ Hash function to the output of Step 6 to produce the final MAC.</td>
</tr>
</tbody>
</table>

C. SHA-1

The Secure Hash Algorithm (SHA) [8] was developed by National Institute of Standards and Technology (NIST) along with NSA.

IV. DESCRIPTION OF NEW ALGORITHMS

This section presents the novel MAJE4 stream cipher algorithm and MACJER-320 algorithm along with JERIM-320 hash function [9] for achieving confidentiality as well as message authentication.

A. MAJE4

a. Main features of MAJE4

1. The encryption sequence can have a large period.
2. The key stream can approximate the properties of a true random stream.
3. MAJE4 is suitable for hardware or software and it uses only primitive computational operations commonly found in microprocessors.
4. It is simple and fast. It uses simple algorithm, which is easy to implement and eases the task of determining the strength of the algorithm.

5. Low memory requirement makes it suitable for handheld type devices with restricted memory.

6. Mixed operators are used for the design of MAJE4. The use of more than one arithmetic and / or Boolean operator complicate cryptanalysis. Primitive operators like + and ^ are used since these operators do not commute and hence cryptanalysis becomes more difficult.

7. It should have a flexible security choice with the key sizes of 128 or 256 bits.

b. Key setup of MAJE4

One can choose either 128 or 256 bit key for the algorithm given in Table III.

128-bit key: The four 32 bit words, ie. key[0], key[1], key[2] and key[3] are considered for storing the key.

256-bit key: The key is stored in eight 32 bit words key[0], key[1], key[2], key[3], key[4], key[5], key[6] and key[7].

TABLE III
ALGORITHM OF MAJE4

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assign the key length kl either as 128-bit or 256-bit.</td>
</tr>
<tr>
<td>2</td>
<td>if kl = 128 then kln = 2, div = 4 else kln = 3, div = 8</td>
</tr>
<tr>
<td>3</td>
<td>if kl = 128 then consider two lsb’s of key[0] and find its decimal equivalent and store in the variable ‘in’. else if kl = 256 then consider three lsb’s of Key[0] and find its decimal equivalent and store in the variable ‘in’.</td>
</tr>
<tr>
<td>4</td>
<td>ran = key[0] ^ key[‘in’]</td>
</tr>
<tr>
<td>5</td>
<td>if kl = 128 then consider two lsb’s of ran and find its decimal equivalent and store in the variable ‘in1’. else if kl = 256 then consider three lsb’s of ran and find its decimal equivalent and store in the variable ‘in1’.</td>
</tr>
<tr>
<td>6</td>
<td>check the 16th bit in ran, if it is 1 then newran = (key[‘in1’] + key[‘in1+1 mod div’]) ^ (key[‘in1+2 mod div’] + key[‘in1+3 mod div’]) else newran = (key[‘in1’] ^ key[‘in1+1 mod div’]) + (key[‘in1+2 mod div’] ^ key[‘in1+3 mod div’])</td>
</tr>
<tr>
<td>7</td>
<td>The output 32-bit word is newran, which can be used to XOR with the corresponding word in the plain text.</td>
</tr>
</tbody>
</table>

b. MACJER-320

The variables used in the MACJER-320 construction are given in Table IV.

TABLE IV
VARIABLES USED IN MACJER-320

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>Shared symmetric key.</td>
</tr>
<tr>
<td>M</td>
<td>Input message.</td>
</tr>
<tr>
<td>B</td>
<td>Number of bits in each block.</td>
</tr>
<tr>
<td>MDA</td>
<td>Message digest algorithm or Hash Function (JERIM-320)</td>
</tr>
<tr>
<td>H</td>
<td>Hash code</td>
</tr>
<tr>
<td>SH</td>
<td>Circular shifted hash code H</td>
</tr>
<tr>
<td>SK</td>
<td>Circular shifted key K</td>
</tr>
</tbody>
</table>

The step-by-step approach of MACJER-320 is given in Table V.

TABLE V
MACJER-320 ALGORITHM

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Make length of K equal to B. Here the initial key K is 320-bit long and the block length B is 512-bit. To make the length of K equal to the block length add as many 0 bits as required to the left of K. Hence add 192, 0 bits to the left of key K.</td>
</tr>
<tr>
<td>2</td>
<td>Prefix and suffix the key along with the message. Divide the key in to two equal parts (256 bits each), and then prefix the message using 256 lsb bits of the key and suffix the message using 256 msb bits of the key.</td>
</tr>
<tr>
<td>3</td>
<td>Apply the message digest algorithm / hash function Now, JERIM-320 is applied to the output of step 2 (i.e. to the combination of the 256 lsb bits of the key, the message, the 256 msb bits of the key) to produce the 320-bit hash code H.</td>
</tr>
</tbody>
</table>

C. JERIM-320

a. Structure of JERIM-320

Fig.1. Shows the outline of the compression function of JERIM-320. It consists of four parallel branches B1, B2, B3 and B4.
b. Single Step Operations

Five rounds are used in JERIM-320 for each 512-bit message block. The sixteen 32-bit sub blocks of the 512-bit block in each round are processed in four parallel branches. The inputs to each single step operations are the sixteen sub blocks, the chaining variables A1, B1,…J1, A2, B2, …J2, A3, B3,….J3, A4, B4,…..J4 of each branch and the constants K [t]. Order of message words, shift values, Boolean functions and constants in each branch and each round are different. There are 16 single step iterations in each round and in all the four branches as shown in Fig. 2. The output of each iteration is copied again into the chaining variables A1, B1,…J1; A2, B2, …J2; A3, B3,….J3; A4, B4,…..J4 and so on.

---

V. SECURITY ANALYSIS

Sections A and B describe an analysis of the stream ciphers RC4 and MAJE4. The security analysis of HMAC is explained in section C. Sections D and E describe an analysis of the properties of Message Authentication Code and Hash Function, which help MACJER-320 to achieve a significantly higher level of security than the popularly used ones.

A. RC4

Some of the published attacks on RC4 are as follows:

1. The first known weaknesses in RC4 were reported in 1995 by Ross [10] and Wagner [11]. They described several classes of keys that have specific weaknesses including predictable output or output that leaks key information. Later a related key attack was observed for long keys (2048 bits) [12].

2. Since the output of RC4 stream cipher is used to encrypt the plain text by bitwise XOR, any observable bias in the output can be used as the basis for an attack. A correlation was detected by Golic [13] between bytes at time t and t+2. Many stronger correlations were later reported by Fluhrer and McGrew [14].

3. Attacks to guess the internal state and then check for consistency with known output have been studied independently by several researchers and the results were published [15]-[17].

4. The most significant attacks on RC4 have been based on exploiting the simplicity of the initialization algorithm to discover an observable bias in the first few bytes of the output sequence. A bias in the second output byte also has been reported [18]. The value zero occurs with twice the expected probability for a random sequence. A bias in the first byte was also reported [19].


6. Klein [20] showed an improved way of attacking RC4 using related keys that does not need the ‘resolved condition’ on the IVs and gets by with a significantly reduced number of frames.

7. Subhamoy Maitra and Goutam Paul gave an independent analysis [21] of Klein’s attack with results similar to our multiple key bytes extension.

8. Vaudenay and Vuagnoux presented a similar attack at SAC2007 [22], which additionally makes use of the fact that the RC4 key is stretched to 256 bytes by repeating it. The same trick was reported by Ohigashi, Ozasa, Fujikawa, Kuwadako and Morii [23], who developed an improved version of the attack.

9. The implication of these findings is that a buffer overflow attack or a similar attack can be used to learn a single state of the generator, which can then be used to predict all random values, such as SSL keys [24]. This type of attack is more severe and more efficient than other known attacks.

These problems with RC4 have raised fears on the security of protocols like the SSL which are using RC4 for providing confidentiality.

B. MAJE4

The MAJE4 is a 128-bit or 256-bit key algorithm and the randomness property of the stream cipher is analyzed by using the five statistical tests like frequency test, serial test, poker test, runs test and autocorrelation test [24]. All the five statistical tests are passed by this generator for all the random streams produced. Hence MAJE4 algorithm can be used very well for encrypting the message of any length.

C. HMAC-SHA-1

1. In HMAC the XOR with Ipad results in flipping one-half of the bits of K. Similarly the XOR with Opad result in flipping the other-half of the bits of K, but a different set of bits. In effect, by passing S and S1 through the compression function of the hash algorithm, we have pseudo randomly generated two keys from K, which add security to HMAC.
2. The recent attacks on Wang et al. and Biham et al. have undermined the confidence in the popular hash functions such as MD5 or SHA-1.

3. As outlined in the paper “keying hash functions for message authentication”, HMACs can be vulnerable to birthday, collision and other attacks [25].

4. Other publications “On the security of HMAC and NMAC based on HAVAL, MD4, MD5, SHA-0 and SHA-1” [26] and “Note on Distinguishing, Forgery and Second Preimage Attacks on HMAC-SHA-1 and a Method to Reduce the Key Entropy of NMAC” [27] have shown how to use the differential distinguishers to devise a forger attack on HMAC.

5. The strongest attack known against HMAC is based on the frequency of collisions for the hash function. With this, HMACs have become more insecure [28].

6. The attack Xiaoyun Wang, Yiqun Lisa Yin and Hongbo Yu recently announced on SHA-1 indicates that the algorithm isn’t quite as strong as it was thought to be, as it takes only 269 steps to find a collision instead of the expected 280 steps.

7. XORing has the effect of randomizing the input almost completely and overcoming any regularity that appear in the output.

E. JERIM-320

1. The main difficulty in cryptanalyzing JERIM-320 comes from the fact that the same message blocks are given as input to each of the four streams in a permuted fashion. The attacker who tries to break JERIM-320 should aim simultaneously at four ways where the message difference passes, which would make the attacks more difficult.

2. By using one message block twice at each single step, it has been made difficult to construct a differential characteristic with high probability.

3. To avoid an attack that depends on brute-force methods, the output from the hash function has been made sufficiently long.

4. While combining the outputs from the four branches, orthogonal operations (+ and ^) are used to create confusion and diffusion which adds to the security.

5. There is a strong avalanche effect; hence a change in a single message bit affects all the registers after five rounds.

6. All shortcut attacks on MD5 target one of the intermediate blocks. Increasing the intermediate value to 320 bits helps to prevent these attacks.

7. The single step operation ensures that changing a small number of bits in the message affects many bits during the various passes. Together with the strong avalanche, helps JERIM-320 to resist attacks similar to Dobbertin’s differential attack [32] on MD4.

It is in this situation that MAJE4 and MACJER-320 have been proposed along with a strong hash function JERIM-320 for achieving better security services.

VI. PERFORMANCE EVALUATION

The performance evaluation is done by comparing RC4 with MAJE4 and MACJER-320 with HMAC-SHA1 as given in section A and section B. The evaluations were done using Pentium IV processor, Linux operating system and C compiler.

A. MAJE4 & RC4

From the timing analysis it can be noted that when we compare RC4 and MAJE4, MAJE4 is almost 1.2 times faster as shown in Table VI. On comparing the memory required for executable files of RC4 and MAJE4, MAJE4 was found consuming lesser space compared to RC4. The memory size required for optimised code for RC4 is 8077 bytes and for MAJE4 is 5435 bytes.

<table>
<thead>
<tr>
<th>TABLE VI</th>
<th>TIMING ANALYSIS &amp; MEMORY REQUIREMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRNGs</td>
<td>MAJE4</td>
</tr>
<tr>
<td>Key length</td>
<td>128-bit</td>
</tr>
<tr>
<td>No. of random numbers generated</td>
<td>1,15,39,399</td>
</tr>
<tr>
<td>No. of random bits per each random number</td>
<td>32</td>
</tr>
<tr>
<td>Total no. of bits produced (speed Mbps)</td>
<td>352.15</td>
</tr>
<tr>
<td>Memory requirement (Bytes)</td>
<td>5435</td>
</tr>
</tbody>
</table>
The total number of operations, memory requirements and the speed performance of MACJER-320 using JERIM-320 hash function and HMAC using SHA-1 hash function, were compared. MACJER-320 produces a MAC of 320 bits whereas HMAC-SHA-1 produces a MAC of 160 bits only. Hence MACJER-320 can definitely provide added security than HMAC-SHA-1.

As shown in Table VII the total number of operations used in MACJER-320 is 3.7 times than that in HMAC-SHA-1. The hash function JERIM-320 in MACJER-320 makes use of four parallel lines of message processing and hence the variables and computations required in JERIM-320 will be more compared to the hash function SHA-1 in HMAC. The multiple operations on the message blocks in MACJER-320 will result in much higher security with a negligible compromise in the speed of operation.

As shown in Table VIII the memory requirement for MACJER-320 is more than that of HMAC-SHA-1. This is due to the increased number of Boolean functions, the need for other operations like add, shift as well as the greater number of lines of message processing used in JERIM-320 than in SHA-1. Even though the speed of MACJER-320 is less than that of HMAC-SHA-1, it is very much within the acceptable limits and hence the advantages due to increase in the security overcomes the reduction in speed.

The SSL designers have chosen to use the then available algorithms RC4 as fast stream cipher and HMAC as hash-based construction for its security services. But few recent findings show that the confidence level in these algorithms is coming down. It is clear that a transition to a newer encryption and message authentication algorithms will be required in the near future, since the information handled is very sensitive. It is in this situation that more secure algorithms MAJE4 and MACJER-320 are suggested which can definitely become good substitutes.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Speed (Mbps)</th>
<th>Memory requirement (Bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACJER-320IM using JERIM-320</td>
<td>13.15</td>
<td>12530</td>
</tr>
<tr>
<td>HMAC using SHA-1</td>
<td>57.58</td>
<td>8074</td>
</tr>
</tbody>
</table>

**VII. CONCLUSION**

The SSL designers have chosen to use the then available algorithms RC4 as fast stream cipher and HMAC as hash-based construction for its security services. But few recent findings show that the confidence level in these algorithms is coming down. It is clear that a transition to a newer encryption and message authentication algorithms will be required in the near future, since the information handled is very sensitive. It is in this situation that more secure algorithms MAJE4 and MACJER-320 are suggested which can definitely become good substitutes.


Sheena Mathew, Reader in Cochin University of Science and Technology (CUSAT), Kochi, Kerala, India has 15 years of teaching experience in Computer Science. She had her graduation from Madurai Kamaraj University and post graduation from the Indian Institute of Science, Bangalore. She is presently a research scholar; her areas of interest being Cryptography and Network Security. She has 8 publications in various international journals and conferences to her credit.

Dr. K. Poulose Jacob, a National Merit Scholar all through, got his degree in Electrical Engineering in 1976 from University of Kerala, followed by his M.Tech. in Digital Electronics and Ph. D. in Computer Engineering from CUSAT, Kochi. He has been teaching at CUSAT since 1980 and currently occupies the position of Professor and Head of the Department of Computer Science. He has served as a Member of the Standing Committee of the UGC on Computer Education and Development. He is on the editorial board of two international journals and has more than 60 papers in various international journals and conferences to his credit. His research interests are in Information Systems Engineering, Intelligent Architectures and Networks.